AMENDMENTS TO THE CLAIMS

Please cancel Claim 2 as indicated below.

Please amend Claims 1 and 3–23 as indicated below.

Please add new Claims 24–27 as indicated below.

1. (Currently amended) A computer-executable method for using machine learning to predict an outcome, the method comprising ~~the steps of~~:

defining a first outcome associated with a first range of medical costs at least as great as a cost threshold;

defining a second outcome associated with a second range of medical costs less than the cost threshold, wherein the second outcome is more likely than the first outcome; and

~~applying a test subset of a data set to~~ processing training data with a machine learning system, wherein said training data is a subset of a data set and is recorded in a computer-readable medium, and wherein the act~~step~~ of ~~applying~~ processing the training data ~~a test subset of data~~ includes:

~~applying~~ selecting a first subset of the training data, the first subset corresponding to [[a]] the first outcome[[,]];

~~applying~~ selecting a second subset of the training data, the second subset corresponding to [[a]] the second outcome and consisting of a set of ~~nearest~~ nearby neighbors to the first outcome[[,]]; and

~~applying~~ selecting a third subset of the training data, the third subset corresponding to the second outcome, wherein the third subset does not consist of nearby neighbors to the first outcome; and

using a plurality of software-based, computer-executable machine learners ~~machine learning methods~~ to develop from the first, second and third subsets one or more sets of computer-executable rules usable to predict the first outcome or the second outcome~~results from the subset~~.

2. Cancelled.

3. (Currently Amended) The method of Claim [[2]]1, wherein the act of selecting the ~~step of applying a~~ third subset ~~of data~~ includes randomly selecting a subset of the ~~test subset of~~ training data corresponding to the second outcome ~~and not consisting of nearest neighbors to the first outcome as the third subset~~.

4. (Currently Amended) The method of Claim 1, wherein the training data ~~step of applying a test subset of data includes applying a test subset of data wherein the test subset of data~~ includes records having an associated medical cost~~outcome variable~~ and a plurality of feature variables.

5. (Currently Amended) The method of Claim 4, further comprising the act ~~step~~ of identifying a ~~set of nearest~~ nearby neighbors by using medical cost values ~~of the outcome variable for the test subset of data~~.

6. (Currently Amended) The method of Claim 5, wherein the act of selecting the ~~step of applying a~~ second subset ~~of data~~ includes randomly selecting a subset of the identified set of nearby~~nearest~~ neighbors as the second subset.

7. (Currently Amended) The method of Claim 5, wherein the act of selecting the ~~step of applying a~~ second subset ~~of data~~ includes selecting as the second subset all of the identified set of nearby~~nearest~~ neighbors.

8. (Currently Amended) The method of Claim 4, further comprising the act~~step~~ of identifying a set of nearby~~nearest~~ neighbors using values of the plurality of feature variables for the training~~test subset of~~ data.

9. (Currently Amended) The method of Claim 1, further comprising the act~~step~~ of validating the one or more sets of rules using the data set.

10. (Currently Amended) The method of Claim 9, wherein the act~~step~~ of validating the one or more sets of rules includes obtaining one or more accuracy measures for the rules using a portion of the data set.

11. (Currently Amended) The method of Claim 9, wherein the act~~step~~ of validating the one or more sets of rules includes obtaining one or more accuracy measures for the rules using the entire data set.

12. (Currently Amended) The method of Claim 11, wherein the act~~step~~ of validating the one or more sets of rules further includes obtaining the one or more accuracy measures for the ~~test subset of the~~training data ~~set~~.

13. (Currently Amended) The method of Claim 11, wherein the actstep of obtaining one or more accuracy measures includes obtaining measures of thea positive predictive value, the a negative predictive value, the a sensitivity, and the a selectivity of the rules.

14. (Currently Amended) The method of Claim 1, wherein the actstep of using a plurality of software-based, computer-executable machine learners machine learning methods includes developing a set of interim rules using the plurality of software-based, computer-executable machine learners machine learning methods, evaluating the set of interim rules, and developing a revised set of interim rules using the results of the evaluating step.

15. (Currently Amended) The method of Claim 14, wherein the actstep of evaluating the set of interim rules includes applying a user-selectable fitness function.

16. (Currently Amended) The method of Claim 14, wherein the actstep of evaluating the set of interim rules includes applying a fitness function based on one or more of the a sensitivity, the a positive predictive value, and the a correlation coefficient of the interim rules.

17. (Currently Amended) A computer-executable method for using machine learning to predict results comprising the actsstepsof:

~~applying~~ processing a representation of a subset of a data set withto a machine learning system, the representation comprising:

first data corresponding to a first outcome, wherein the first outcome is associated with a range of medical costs at least as great as a predetermined threshold amount;

second data corresponding to a second outcome, wherein the second outcome is associated with a range of medical costs lower than the predetermined threshold amount, wherein the second data consists of a set of nearby neighbors to the first outcome, and wherein the second outcome is less likely than the first outcome; and

third data corresponding to the second outcome, wherein the third data is different than the second data;

repeating for a plurality of cycles:

using a plurality of software-based, computer-executable machine learnersmachine learning methods to develop a set of computer executable rules from the ~~applied~~ processed representation of the subset of the data set;

evaluating the set of computer-executable rules using a user-selectable fitness function; and

modifying the machine learning methods executed by a plurality of software-based, computer-executable machine learners by using the results of the evaluating act~~step~~; and

presenting a final set of computer-executable rules usable to predict the first outcome or the second outcome.

18. (Currently Amended) The method of Claim 17, wherein the act~~step~~ of evaluating a set of rules includes using a user-selectable fitness function based on one or more of: ~~the~~ a number of true positives, ~~the~~ a number of true negatives, ~~the~~ a number of false positives, and ~~the~~ a number of false negatives that the set of rules obtains from the subset of the data set.

19. (Currently Amended) The method of Claim 17, wherein the actstep of evaluating a set of rules includes using a user-selectable fitness function based on a the sensitivity and a the positive predictive value of the rules.

20. (Currently Amended) The method of Claim 17, wherein the actstep of evaluating a set of rules includes using a user-selectable fitness function based on a the sensitivity, a the positive predictive value, and a the correlation coefficient of the rules.

21. (Currently Amended) The method of Claim 17, further comprising, in at least one of the plurality of cycles, developing one or more new representations of the data for use by the plurality of software-based, computer-executable machine learners machine learning methods in a subsequent cycle.

22. (Currently Amended) A computer-executable method for using machine learning to predict a positive or a negative outcome, where the positive outcome is less likely than the negative outcome, the method comprising ~~the steps of~~:

defining a positive outcome associated with a range of medical costs equal to or greater than a cost threshold;

defining a negative outcome associated with a range of medical costs less than the cost threshold; and

~~applying a test subset of a data set to~~ processing training data with a machine learning system, wherein said training data is a subset of a data set and is recorded in a computer-readable medium, and wherein the act~~step~~ of ~~applying~~ processing the training data ~~a test subset of data~~ includes:

~~applying~~ selecting a first subset of the training data, the first subset corresponding to ~~a first~~ the positive outcome[[,]];

~~applying~~ selecting a second subset of the training data, the second subset corresponding to ~~a second~~ the negative outcome and consisting of a set of nearest neighbors to the ~~first~~ positive outcome;~~, and~~

~~applying~~ selecting a third subset of the training data, the third subset corresponding to the ~~second~~ negative outcome, wherein the third subset does not consist of nearest neighbors to the positive outcome; and

using a plurality of software-based, computer-executable machine learners ~~machine learning methods~~ to develop from the first, second and third subsets of the training data one or more sets of computer-executable rules usable to predict ~~an~~ either the positive outcome or the negative outcome ~~from the subset~~.

23. (Currently Amended) The method of Claim 22, wherein the act~~step~~ of using a plurality of software-based, computer-executable machine learners ~~machine learning methods~~ to develop one or more sets of rules includes applying a user-selectable fitness function to develop the one or more sets of rules.

24. (New) The method of Claim 22, wherein the negative outcome is at least thirty times more likely than the positive outcome.

25. (New) The method of Claim 1, wherein the plurality of software-based, computer-executable machine learners executes a neural network machine learning process.

26. (New) The method of Claim 1, wherein the plurality of software-based, computer-executable machine learners executes a decision tree machine learning process.

27. (New) The method of Claim 1, wherein the first, second and third subsets each include approximately equal amounts of data.